

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE Ontological Considerations for Uncertainty Propagation in High Level Information Fusion				5a. CONTRACT NUMBER W911NF-11-1-0176	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 611102	
6. AUTHORS Mark Locher, Paulo Costa				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES George Mason University Office of Sponsored Programs 4400 University Drive, MSN 4C6 Fairfax, VA 22030 -4422				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211				10. SPONSOR/MONITOR'S ACRONYM(S) ARO	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) 59841-MA.7	
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT Uncertainty propagation in a level 2 high level information fusion (HLIF) process is affected by a number of considerations. These include the varying complexities of the various types of level 2 HLIF. Five different types are identified, ranging from simple entity attribute refinement using situation status data to the development of a complete situation assessment assembled from applicable situational fragment data. Additional considerations include uncertainty handling in the input data, uncertainty representation, the effects of the reasoning technique used in the fusion process, and output considerations. Input data considerations include the data's relevance to the					
15. SUBJECT TERMS High level fusion, input uncertainty, process uncertainty, output uncertainty, uncertainty representation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Paulo Costa
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 703-993-9989

Report Title

Ontological Considerations for Uncertainty Propagation in High Level Information Fusion

ABSTRACT

Uncertainty propagation in a level 2 high level information fusion (HLIF) process is affected by a number of considerations. These include the varying complexities of the various types of level 2 HLIF. Five different types are identified, ranging from simple entity attribute refinement using situation status data to the development of a complete situation assessment assembled from applicable situational fragment data. Additional considerations include uncertainty handling in the input data, uncertainty representation, the effects of the reasoning technique used in the fusion process, and output considerations. Input data considerations include the data's relevance to the situation, its credibility, and its force or weight. Uncertainty representation concerns follow the uncertainty ontology developed by the W3C Incubator Group on Uncertainty Reasoning. For uncertainty effects of the fusion process, a basic fusion process model is presented, showing the impacts of uncertainty in four areas. Finally, for output uncertainty, the significance of a closed-world versus open-world assumption is discussed.

TECHNICAL REPORT TR201312A

Ontological Considerations for Uncertainty Propagation in High Level Information Fusion

Mark Locher
George Mason University and SRA, International
Fairfax VA USA
mlocher@gmu.edu

Paulo C. G. Costa
George Mason University
Fairfax VA USA
pcosta@gmu.edu

Abstract— Uncertainty propagation in a level 2 high level information fusion (HLIF) process is affected by a number of considerations. These include the varying complexities of the various types of level 2 HLIF. Five different types are identified, ranging from simple entity attribute refinement using situation status data to the development of a complete situation assessment assembled from applicable situational fragment data. Additional considerations include uncertainty handling in the input data, uncertainty representation, the effects of the reasoning technique used in the fusion process, and output considerations. Input data considerations include the data's relevance to the situation, its credibility, and its force or weight. Uncertainty representation concerns follow the uncertainty ontology developed by the W3C Incubator Group on Uncertainty Reasoning. For uncertainty effects of the fusion process, a basic fusion process model is presented, showing the impacts of uncertainty in four areas. Finally, for output uncertainty, the significance of a closed-world versus open-world assumption is discussed.

Keywords - High level fusion, input uncertainty, process uncertainty, output uncertainty, uncertainty representation

I. INTRODUCTION

The past 20 years have seen an explosion of systems and techniques for collecting, storing and managing large and diverse sets of data of interest to a number of communities. These data are collected by a wide variety of mechanisms, each of which has varying considerations that influence the uncertainty in the data. In order to provide useful information for a particular question or problem, the relevant data ("evidence") must be identified, extracted and then fused to provide insight or answers to the question or problem. The information fusion community has developed a widely accepted functional layered model of information fusion. These layers can be divided into low level and high-level fusion. At all levels, the data going into a fusion process is recognized as having uncertainty, which affects in various ways the degree of certainty in the output of the process. Low-level fusion has been widely explored, primarily through the radar tracking community, and issues of uncertainty determination and propagation are well understood [1].

High-level fusion, on the other hand, requires reasoning about complex situations, with a diversity of entities and various relationships within and between those entities. This reasoning is often expressed symbolically, using logic-based approaches [2]. There has been significant work in using ontological approaches in developing fusion techniques and

some of these approaches have taken uncertainty considerations into account (e.g. [3] [4] [5] [6]). Various techniques exist to model and propagate uncertainty in a fusion process, with varying strengths and difficulties. This suggests that their relative performance in a fusion system should vary significantly depending on the types and nature of the uncertainties within both the input data and the context of the problem set modeled with the fusion system. Unfortunately, there is no consensus within the fusion community on how to evaluate the relative effectiveness of each technique. Work in this area will be hampered until the evaluation question is at least better defined, if not resolved.

The International Society for Information Fusion (ISIF) chartered the Evaluation of Technologies for Uncertainty Reasoning Working Group (ETURWG) to provide a forum to collectively address this common need in the ISIF community, coordinate with researchers in the area, and evaluate techniques for assessing, managing, and reducing uncertainty [7]. In its first year, ETURWG defined its scope and developed the uncertainty representation and reasoning evaluation framework (URREF) ontology. The URREF ontology aims to provide guidance for defining the actual concepts and criteria that together comprise the comprehensive uncertainty evaluation framework [8]. It is evident that part of the issue in evaluating different uncertainty representation systems is to properly understand how a high-level fusion process works and how uncertainty is propagated through the process.

This paper aims to help establish the various considerations about how uncertainty affects a HLIF process. It will begin by defining what is meant by a HLIF process, and then focus on one class of HLIF, the level 2 HLIF. From there, it will define a taxonomy of Level 2 HLIF, where increasing complexity of level 2 HLIF types have additional uncertainty considerations. Then it explores uncertainty propagation issues associated with uncertainty in the input data, the uncertainty effects of both the fusion reasoning process and the representation scheme, and the output uncertainty. It concludes with a top-level discussion of a mathematical approach applicable to these considerations.

II. DEFINITION OF HIGH-LEVEL FUSION

A widely accepted definition of High-Level Information Fusion (HLIF) is that it refers to the fusion processes classified as level 2 and above within the revised Joint Directors of Laboratories data fusion model. This model establishes five functional levels, as defined in [9] and repeated in Table 1 below.

Table 1: JDL Fusion Levels [9]

Level	Title: Definition
0	Signal / Feature Assessment: Estimate signal or feature state. May be patterns that are inferred from observations or measurements, and may be static or dynamic, and may have locatable or causal origins
1	Entity Assessment: Estimation of entity parametric and attributive states (i.e. of individual entities)
2	Situation Assessment: Estimate structures of parts of reality (i.e. of sets of relationships among entities and implications for states of related entities.)
3	Impact Assessment: Estimate utility/cost of signal, entity or situation states, including predicted utility / cost given a system's alternative courses of action
4	Process Assessment: A system's self-estimate of its performance as compared to desired states and measures of effectiveness.

A key item is that these assessments are not just a combination of information, but they are also analytic judgments. For example, a level 2 fusion process is more than a unified display of information (e.g. a common operational picture); rather, it requires explicit statements about how certain specific elements of reality are structured, in order to address specific questions that a user of that process wants answered. Level 2 fusion essentially answers the question “what is going on?” Level 3 fusion addresses “what happens if ...?”, where “if” is followed by a possible action or activity (level 3 is often predictive). Level 4 involves steering the fusion system, including adjusting data collection based on an assessment of already-collected data. There has been some discussion regarding the boundary between level 1 and level 2. Das, for instance, considers identification and object classification as beyond level 1, suggesting that this type of fusion should be a level 1+ [10]. Steinberg, on the other hand, considers this to be clearly level 1 [9]. Sowa's ontological categories provide insight into this question, and can be used to illuminate some factors on uncertainty propagations considerations. In the present work, these ontological categories were used as a basis for defining a taxonomy of level 2 fusion.

III. TAXONOMY OF LEVEL 2 HLIF

Sowa defined twelve ontological categories, and together they comprise a very attractive framework for analyzing fusion processes at level 2. He suggests that one way of categorizing entities in the world is to consider them from three orthogonal aspects [11]. The first is whether they are physically existing or abstract. Abstract entities are those that have information content only, without a physical structure. This includes the idea of geometric forms or canonical structures (e.g. idea of a circle), or entities like computer program source code.

The second aspect defining the ontological categorization is whether the entity is a continuant (i.e., having time-stable recognizable characteristics) or an occurrent (i.e., significantly changing over time). This means that an entity can either be an object (a continuant) or a process (an occurrent – also called an event). The third and final aspect of his ontological categorization is the degree of interrelatedness with other

objects and processes. At the independent level, an entity is considered by itself, without reference to other entities. At the relative level, an entity is considered in single relation to another entity. Finally, the idea of mediating takes into account two items: the number and complexity of the various interrelationships among the entities, and the unifying idea – its purpose or reason – that allows one to define a situation or a structure that encompasses the relevant entities [11].

The combination of these three aspects results in the 12 ontological categories shown in Table 2. Table 3 provides a more detailed definition of each ontological category and provides some examples.

A key point in looking at this ontological categorization is that one must understand the context and viewpoint from which a given entity is categorized, and that changes to either of these two might result in different categorizations for the same entity. To illustrate this point, an airplane can be considered as either an independent object flying in the air, or a complex mediating structure with thousands of component objects and processes that work together for the purpose of achieving aerial flight. The viewpoint one takes depends on the context one is interested in. In the airplane example, it depends on whether one is tracking a particular aircraft using a variety of sensors, or attempting to determine the various capabilities of a new aircraft type.

Table 2: Sowa's Categories [11]

	Physical		Abstract	
	Continuant	Occurrent	Continuant	Occurrent
Independent	Object	Process	Schema	Script
Relative	Juncture	Participation	Description	History
Mediating	Structure	Situation	Reason	Purpose

It is tempting to suggest that Sowa's three relationship levels correspond to the JDL levels 1 / 2 / 3 (i.e., Independent, Relative, and Mediating, respectively). However, this has at least three major problems. First, Sowa's relative level is focused on a single relationship between two entities, while JDL level 2 can (but does not have to) consider multiple relationships in and between multiple entities. Second, JDL level 2 situation assessment includes making assessments about the purpose or reason for the situation. This reason or purpose is the key characteristic that distinguishes one situation from another. A raucous sports team victory celebration, a protest and a riot share many entities and relationships, but understanding the reason/purpose behind it can make a significant difference to a chief of police. Third, there are level 1 inferences that depend on the existence of fixed relationships between entities.

To illustrate the latter point above, consider the case of an intercepted radar signal that has been classified as having come from a specific type of radar system. Now let us suppose that the radar type is tightly associated with a larger system, such as the AN/APG-63 radar on older versions of the US F-15 aircraft [12]. If one has detected the APG-63 radar, one also has very high confidence that one has detected an F-15 aircraft. This F-15 object identification occurs because there is a fixed relationship between the two objects (it's not a 100% relationship, as the APG-63 is also installed on fourteen United

States Customs and Border Protection aircraft [13]). This situation is a clear example of a fixed relationship between entities (i.e., AN/APG-63 used in F-15 fighters) that supports a level 1 object identification, thus making it applicable to directly associate JDL level 1 to Sowa's Independent relationship.

Table 3: Definitions [11]

	Definition	Examples
Object	Any physical continuant considered in isolation	Any specific existing item (e.g. car serial number 123, etc.)
Process	The changes that occur to an object over time, with a focus on the changes	Explosion, most action verbs
Schema	The form of an continuant	Circle, language concepts for classes of objects (e.g. cat, airplane)
Script	The time or time-like sequence of an occurrent	Process instructions, software source code, radar track file
Juncture	Time-stable relationship between two objects	Joint between two bones, connection between parts of a car
Participation	Time-varying relationship between two objects, or a process related to an object	Artillery firing a shell, radio communication between two people
Description	An abstraction about the types of relationships that can exist between continuants	The idea behind concepts like "join", "separate", "works for", "mother of", etc.
History	The recorded information about an occurrence as it relates to one or more continuants	Video file of a traffic intersection
Structure	A complex continuant with multiple sub-continuant and many relationships. Focus is on the stability of the continuant	Composition of an army, layout of a chemical plant
Situation	A complex occurrent with multiple continuants and many relationships. Focus is on the time sequence of changes among the objects and processes	A birthday party, road traffic in a metropolitan area
Reason	The intention behind a structure	Differentiates a chemical weapon factory from a fertilizer factory
Purpose	The intention driving a situation	Intention that differentiates going to war from conducting a military exercise

Now consider the case where the radar is associated with a Surface-to-Air Missile (SAM) system, such as the Tin Shield acquisition radar and the SA-10 Grumble SAM system. The SA-10 system consists of multiple separate vehicles, not a single vehicle. The radar vehicle is physically separate from the other vehicles. It is possible for the Tin Shield radar to be used as a stand-alone search radar [14]. In this case, detection of the Tin Shield radar signal may indicate the presence of the SA-10, but it may not.

A key differentiator between JDL levels 1 and 2 is the focus on an object versus on multiple objects in relationship to each other. Yet, as illustrated by the two later examples, a JDL level 1 assessment can use techniques that are grounded in Sowa's relative level. In general, determining an object's level 1 attributes and states often depends on fusing different sensor outputs of processes that an object has undergone – thus making use of participation level information.

Using Sowa's categories, one can create the taxonomy of level 2 situations shown in Figure 1. This taxonomy ranges widely in complexity and analytic inferences required. There are five cases presented in the Figure, each created by first determining whether one is dealing with a known situation, or whether the situation itself must be inferred. In general, the least complex case is for known situations where one is determining / refining the attribute of an entity. This case straddles the level 1 / 2 line. It is object / process identification where the relationship between elements within the object of interest may vary. An example is the radar / vehicle case above. The defined situation is that a Tin Shield radar has been detected at a particular location. The question is whether an SA-10 battery (a higher level object) is at that location, or whether the radar is operating in a stand-alone mode (whether operationally, for system testing, or for system maintenance). The inferences generally are based on schema-based evidential reasoning (e.g. "there is a 95% chance that this radar will be associated with an SA-10 battery in its immediate vicinity").

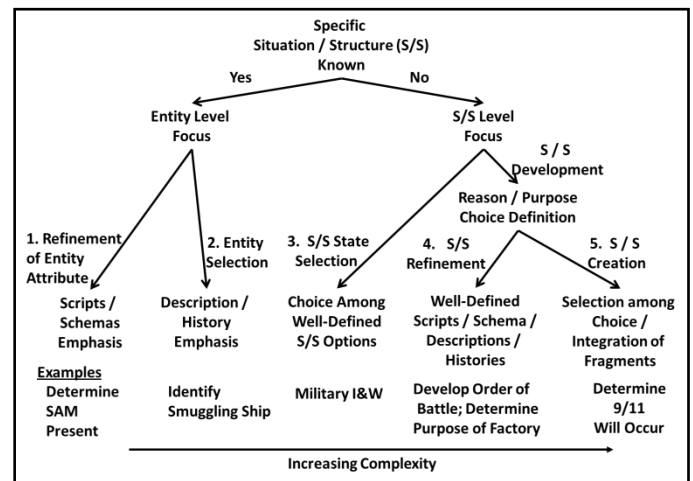


Figure 1: Types of Situation Assessments

The second case is a step up in complexity, where the situation is well defined but the objective is to identify a specific object of interest within the situation. For example, one might have very credible evidence that a terrorist group will attempt to smuggle a radiological bomb into the United States via a freighter. In this case, the situation itself is known (one knows the purpose / intention), but the actors may be hidden. Inferring which freighter (an object identification) is a likely carrier of the bomb is the question of interest. Another example would be to determine who committed a robbery of a bank, when one has a video of the act itself (the situation is a robbery). In this case, the evidence is extracted from a variety

of sources, which can be classified as being junctures, participations, histories or descriptions.

The inferential process generally becomes more complex when the specific situation itself is not known, but must be inferred. The taxonomy outlines three such cases, each with an increasing level of complexity. The first is when the specific situation is not known, but there is a set of well-defined situation choices to select from. This case is a situation version of a state transition. A classic example is the military indications and warning question, which can be raised when an increase in activity at military locations in a country is detected. The question then becomes “what is the purpose of the activity?” Four major choices exist: a major military exercise, suppression of domestic unrest, a coup d' etat, or preparing to go to war. Each is a relatively well-defined situation with known entities, attributes and relationships. The selection among them becomes a pattern-matching exercise.

The next level of complexity occurs when not only is the situation itself unknown, the situation itself must be developed. Unlike the case above, the issue now is not choosing among a set of possible situations but to build the situation from the data. This case can be divided into two subcases. In the first subcase, one has a series of templates that can be used in developing aspects of the situation. For example, in developing an enemy order of battle for a standing nation-state's military, one has a basic understanding of the objects and relationships that constitute a modern military force. A country may not have all of the elements, and the organizational structure will vary. Yet, it is very likely that the structure and deployment will follow patterns similar to those used by other countries.

The second subcase is the most complex situation. Here, one must develop a situation where the basic purpose itself must be determined. For example, consider the case when a government agency is notified that something is significantly amiss, with enough information to spark interest, but not enough to understand what is happening. In that case, the evidence must be assembled without a common template to guide the fusion. Rather, the evidence must be fused using fragmentary templates, that themselves must be integrated to provide the overall situation. Integrating the data to “connect the dots” that could have predicted the September 11, 2001 commercial airliner strikes on the World Trade Center and the Pentagon falls into this category. Note also that this case also straddles the level 2 / level 3 fusion line, since determining the purpose in this case has a predictive element with possible courses of actions and outcomes.

IV. UNCERTAINTY PROPAGATION IN HLIF

In any fusion process, one follows a fundamental reasoning process, which logically uses a series of reasoning steps, often of an “if, then” form. Beginning with a set of events, we form a chain of reasoning to come to one or more conclusions. Figure 2a models a simple case, while Figure 2b gives an example of that case. More complex structures can be easily created [15].

The ETURWG found that within this fundamental process there were at least four areas for uncertainty considerations: the uncertainty in the input data, the uncertainty associated with

representation within the fusion system, the uncertainty effects of the reasoning process, and the resultant uncertainty in the outputs of the process [7, 8]. The subsections below address some of the ontological considerations associated with the first three factors. Issues associated with output uncertainty are treated in section V.

A. Uncertainty in the Input Data

All conclusions are ultimately grounded on evidence, drawn from a variety of data sources. But often evidence is “inconclusive, ambiguous, incomplete, unreliable, and dissonant.” Any conclusions drawn from a body of evidence is necessarily uncertain. Schum [15] found that one must establish the credentials of any evidence used in a reasoning process. These credentials are its relevance to the question / issue at hand, its credibility, and its weight or force [16]. This suggests that one should elaborate on the fundamental reasoning process from Figure 2 with the additional items shown in Figure 3.

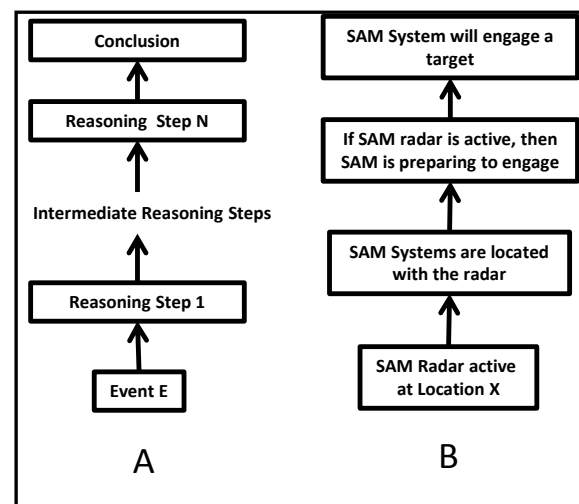


Figure 2: Fundamental Reasoning Process

Data becomes evidence only when it is relevant. Relevance assesses whether the evidence at hand is germane to the question(s) being considered. Irrelevant information makes no contribution to the conclusion drawn, and potentially confuses the fusion process by introducing extra noise. Evidence can be either positively (supportive) or negatively (disconfirmatory) relevant to a particular hypothesis. Any analytic effort is obliged to seek and evaluate all relevant data.

Once data is shown to be relevant to a particular problem (i.e., it becomes evidence), Schum points out that there is an important but often overlooked distinction between an event (an object, process, juncture or participation in Sowa's ontological categories) and the evidence about that event or state. That is, Joe's statement “I saw Bob hit Bill with a club” does not mean that such event actually happened, and should be seen only as evidence about it. Credibility establishes how believable a piece of evidence is about the event it reports on. Schum identified three elements of credibility [17]; the ETURWG added self-report as a distinct element (see Table 4 for elements and definitions) [7].

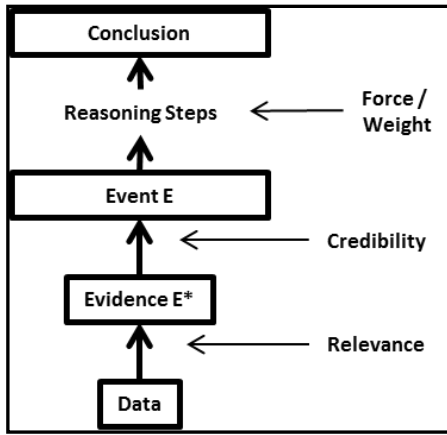


Figure 3: Evidential Factors

Table 4: Elements of Evidential Credibility

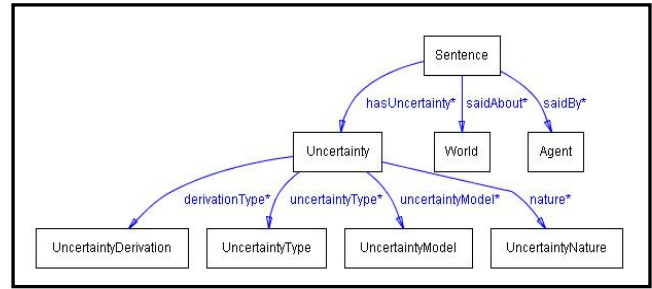
Veracity: Source is telling what it believes to be true (note that the source may be deceived)
Objectivity: Source has received the evidence on which it based its reporting. This includes consideration of system biases and false alarms
Observational Sensitivity: Source has the ability to actually observe what it reports (e.g. Observer actually has the visual acuity needed to see what was going on, or an electronic intercept was of such low quality the operator guessed part of the conversation)
Self-Report: Source provides a measure of its certainty in its report (e.g. a human source hedges her report with “it’s possible that...” or a sensor reports that detection was done at a signal to noise ratio of 4)

The force (or weight) of the event establishes how important the existence of that event is to the conclusion one is trying to establish. By itself, the event “Bob hit Bill with a club” would have a significant force in establishing a conclusion that Bill was seriously injured. It would have less force in establishing that Bill was committing a violent act and needed to be stopped at Bill, and even less force in concluding that Bob was angry at Bill. Figure 3 shows that credibility can have an effect on the force of an event on the conclusion. For example, if the credibility of Joe’s testimony about Bob hitting Bill with a club is low, the certainty of a conclusion that Bob’s hitting was the cause of Bill’s injuries would be less than if Joe testimony’s credibility was high. Schum investigated a number of different ways in which considerations about data credibility could affect the overall conclusions. One of his most interesting findings is that, under certain circumstances, having credible data on the credibility of a data source can have a more significant force on the conclusion than the force of the event reported in the data [15].

B. Uncertainty in the Representation

Uncertainty varies in its forms and manifestations. Therefore, the uncertainty representation scheme used has an effect on what can or cannot be expressed. To see this, one first needs to have an understanding on the different types of uncertainty. The W3C Incubator Group exploring uncertainty reasoning issues for the World Wide Web developed an initial ontology of uncertainty concepts, shown in Figure 4 [18].

Figure 4: Uncertainty Ontology



A Sentence is a logical expression in some language that evaluates to a truth-value (formula, axiom, assertion). For our purposes, information will be presented in the form of sentences. The World is the context / situation about which the Sentence is said. The Agent represents the entity making the Sentence (human, computer etc.). Uncertainty is associated with each sentence, and has four categories. Three of those are described in Table 5, along with their significance for uncertainty propagation in a HLIF process.

Table 5: Definition of Uncertainty Categories

<u>Uncertainty Derivation</u>
Objective: Derived in a formal way, repeatable derivation process. <u>Significance</u> - level of uncertainty can be reliably estimated
Subjective: Judgment, possibly a guess. <u>Significance</u> - Level of uncertainty may be unpredictable
<u>Uncertainty Nature</u>
Aleatory: Uncertainty inherent in the world <u>Significance</u> - Additional data will not resolve uncertainty
Epistemic: Uncertainty in an agent due to lack of knowledge <u>Significance</u> - Uncertainty could be resolved by additional evidence gathering, which eliminates the lack of knowledge
<u>Uncertainty Type</u>
Ambiguity: Referents of terms are not clearly specified <u>Significance</u> - The same evidence may not distinguish between two or more possibilities
Empirical : Sentence about a world is either satisfied or not satisfied in each world, but it is not known in which worlds it is satisfied; this can be resolved by obtaining additional information (e.g., an experiment) <u>Significance</u> - Uncertainty can be resolved with additional information
Randomness (Type of empirical uncertainty): sentence is an instance of a class for which there is a statistical law governing whether instances are satisfied <u>Significance</u> - The empirical uncertainty has a predictable basis for making an estimate, using the appropriate statistical law
Vagueness: No precise correspondence between terms in the sentence and referents in the world <u>Significance</u> - Uncertainty due to a lack of precision
Incompleteness: information about the world is incomplete / missing <u>Significance</u> - Uncertainty increases because assumptions / estimates of information must be used, rather than the actual information. May not have a basis for making an estimate
Inconsistency: no world can satisfy the statement. <u>Significance</u> - Data is contradictory; must resolve source of contradiction (Can occur when deception is used)

The last category in the ontology is Uncertainty Model, capturing the various approaches that can be used to model uncertainty in a reasoning process. These include (but are not limited to):

- Bayesian Probability Theory
- Dempster-Shaffer Evidence Theory
- Possibility Theory
- Imprecise Probability approaches
- Random Set Theory
- Fuzzy Theory / Rough Sets
- Interval Theory
- Uncertainty Factors

A critical item in uncertainty propagation is the proper fit between the types of uncertainty in the input data and in the model(s) used in the fusion reasoning process. Failure to account for all of the uncertainty types in the input data can result in an erroneous process output. A classic survey of uncertainty models, with a discussion on applicable uncertainty types, is given in [19], with a recent review of the state-of-the-art in [20]

C. Uncertainty in the HLIF Fusion Process

To explore the ontological considerations of the uncertainty propagation in a HLIF fusion process, we need to have a basic fusion process model. We will concentrate on the level 2 fusion process only, and leave out significant detail on the processes at the other levels. Figure 5 shows this model. The first thing to observe is that the raw data can come in at any level, as evidenced by the incoming arrows at the right side of the figure. The model does not require that all data be signal or feature (Level 0) data, which is then aggregated into higher-level conclusions. For instance, object identification data (level 1) could come from an on-scene observer or from an image analyst reporting on an image. Communications intercepts or human reporting could provide evidence on relationships (level 2) or future intentions (level 3). Note that if a level 3 fusion process is active, its outputs could affect the level 2 process in two places. It can either be a controlling variable in the fusion process itself, or it can affect the interpretation and extraction of evidence. However, a level 3 process will have an effect only if it has separate evidence that is not being used in the level 2 fusion process (otherwise one has circular reporting).

There are four basic processes in this model. The first is the **fusion** process itself, which is usually some form of a model-based process. These models most often take the form of Bayesian networks [10, 21, 22], although alternative approaches have been proposed using graphical belief models [23] and general purpose graphical modeling using a variety of uncertainty techniques [14].

Another important aspect of this model that must be emphasized is that not all of the evidence that goes into the model-based process is (or is assumed to be) in an immediately usable form. Some data must have the appropriate evidence extracted from it. This is where the uncertainty considerations associated with representation within the fusion system come into play. For example, the raw level 2 data may be a series of

people association data, which must be combined into a social network analysis to reveal the full extent of the relationships.

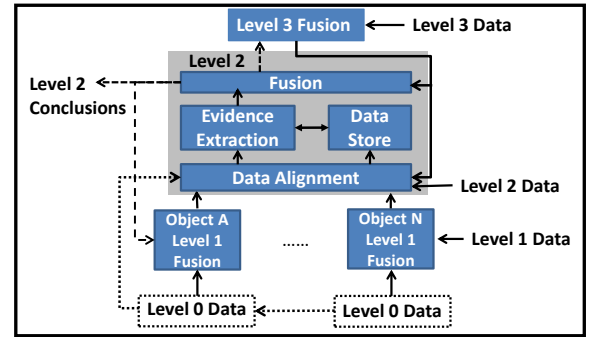


Figure 5: Level 2 Fusion Process Model

Another example may be that one is interested in whether two ships met and transferred cargo in the open ocean. Suppose that you have a track file on each ship which has long revisit rates between collections. This does not provide an obvious indication that the ships met and stopped for a while. But the track files show that both ships were on tracks that did put them at a common location at a given period, and that the average speed dropped significantly during the time a meeting could have occurred (implying that the ships may have stopped for a while). Given this data, one could conclude with some level of certainty that they did meet and stopped to transfer something. This level of certainty is driven by at least two factors: the quality of the track file data (establishing how certain one is in concluding that the tracks allowed them to meet), and how likely is it that two ships showing these track characteristics actually would have met and stopped.

A significant part of the **evidence extraction** process could be comparison to historical or reference data. For example, a vehicle may be moving outside of a normal shipping lane / airway or off-road. This requires a reference to a map base. For this reason, the process model includes a **data store**, for both reference information and for previous data.

The last part of the model is a **data alignment** process. Data may come in with different reference bases, and need to be aligned to a common baseline in order to be used in the extraction and fusion processes.

Finally, note that the level 2 process includes the possibility of a direct use of level 0 data. An area of active research is the multi-source integration of level 0 data that is not of sufficient quality, or that does not have enough quantity to allow a high quality single-source conclusion.

V. MATHEMATICAL CONSTRUCT

A. Model

Several authors have developed mathematical constructs for use in assessing the uncertainty of a situation assessment [2, 25]. Our model is a version of the one put forth by Karlsson [26], modified using the terminology put forth by Franconi [27]. Karlsson's version focuses only on relationships, and does not explicitly include predicates and attributes. While one can model predicates and attributes using relationships, it is

cleaner to separate the entity space from the attribute space. In addition, the construct formed in this paper acknowledges level 2 HLIF as explicitly including entity attributes as well as relationships between entities. Including attributes as separate from entity relationships, rather than defining relationships to include attribute states makes this clearer. Per [27], the language consists of:

E_n , the 1-ary predicates

A_k , the attributes (stated as 2-ary predicates)

R_p , n-ary predicates for all relationships

There is an interpretation function $I = \langle D, \cdot \rangle$ where domain D is a non-empty set $= \Omega \cup B$, Ω is the set of all entities, B is the set of all attribute values and $\Omega \cap B = \emptyset$. Then

$$E_i^I \subseteq \Omega$$

$$A_i^I \subseteq \Omega \times B$$

$$R_i^I \subseteq \Omega \times \Omega \times \dots \times \Omega = \Omega^n$$

x_i are the specific instances and $x_i \in \Omega$

We can make at least three uncertainty assessments. For any specific entity tuple (x_1, \dots, x_n) , we have a level of uncertainty as to whether that tuple is a member of a specific relationship. For a generic uncertainty measure u_T , the basic equation for whether a tuple is correctly associated with a defined relationship is

$$u_{Tj}((x_1, \dots, x_n)_j \in R_j \mid E_B, S, I) \quad (1)$$

where E_B is the body of evidence used in making the assignment, and S, I are any already known situation or impact states. A similar equation holds for attribute uncertainty.

We can also have uncertainty as to whether a relationship that we see in the data is the relationship of interest. Given a set of k possible relationship and a body of evidence E_B for a particular relationship R_{current} , we can assess the following uncertainty:

$$u_{Rk}((R_{\text{current}} = R_k \mid E_B, S, I) \quad (2)$$

Again, a similar uncertainty equation holds for attribute uncertainty. Situation assessment depends on the relationships in the situation. A situation then can be defined as

$$S \stackrel{\text{def}}{=} (R_1, \dots, R_k, A_1, \dots, A_n) \quad (3)$$

Finally, we have an uncertainty measure u_s . Given a set of m possible situations and a body of evidence E_B for a particular relationship S_{current} , we can assess the following uncertainty:

$$u_s(S_{\text{current}} = S_m \mid E_B, I) \quad (4)$$

In addition to uncertainties in the evidence and in the reasoning process, equation (4) also allows us to account for uncertainties in the situation definition. Equation 3 implies that every situation can be precisely defined as a set of specific relationships and attributes. But what if a relationship or attribute is missing in a particular situation instance? For example, a canonical birthday celebration in the United States includes a cake with a number of lit candles on it. If there are

no candles on the cake, does this mean it is not a birthday celebration?

B. Application to Situation Assessment Taxonomy

We can use this model to better understand the varying complexities of the different situation assessment cases given in section 3. For the simplest case, entity attribute refinement, we see that we have a very simple situation (“emitter operational in the environment”). From the existence of one object (the Tin Shield radar), we are inferring the existence of a second object (the SA-10 SAM system). This is a binary relation, based on a Sowa Juncture (x_1, x_2) . With this binary relation, we are operating with a single instance of equation (1). We only have the uncertainty measure for “Tin Shield” and “SA-10” to be in juncture. For the second case, entity selection, we again have a defined situation, but now are seeking a specific object within multiple choices of objects. We are operating at the level of equation (2) – we are seeking a specific relation that ship i is the ship of interest. Based on the evidence, we will create multiple tuples for the different relationships that could lead us to the ship (using equation (1)) and then combine the results to get to equation (2).

For the third case, structure / situation selection, we invoke equation (4) as the basic equation. We are choosing between multiple choices as to what the situation is. We use equation (1) to determine if various relationships exist, and based on those findings, determine which situation model is the correct one for this body of evidence. For the fourth case, structure / situation refinement, we again use equations (1) and (4). But we also use equation (2) to determine what the exact set of relationships is. Case 4 differs from case 3 in that we are trying to determine what the relationships are that are appropriate for this situation (or structure).

For the fifth case, structure / situation creation, we have all of the uncertainties addressed above, and we add an uncertainty not immediately obvious in the generic equations. Relook equation (4). One of the stated requirements is that we are selecting among a set of defined situations. This essentially is a closed world assumption. However, in case 5 we are building the situation, rather than determining which situation among a choice of situations is the applicable one. We still have a number of models to choose from, but they are more fragmentary than in previous cases. The previous cases represent more of a “pieces of the puzzle” approach, where one is assembling the puzzle according to one or more available pictures to help guide you. Case 5 represents the case where we one is assembling the puzzle without a picture or set of pictures to guide one. Rather, you are assembling the puzzle guided by basic puzzle rules about matching shapes and picture colors. So, in case 5, we are also determining what the applicable S_k s are.

VI. DISCUSSION

Up to this point we have been able to attest the existence of a number of uncertainty propagation considerations when analyzing a level 2 HLIF. Most of these are not necessarily obvious at a first glance, which suggests the importance of a framework that supports the analytical process. The framework proposed in this paper is meant for supporting the

The diagram illustrates a multi-level processing framework. At the bottom, 'Data' is input into 'Data Alignment' (Level 1). 'Data Alignment' also receives 'Data Uncertainty (relevance, credibility)'. 'Data Alignment' outputs to both 'Evidence Extraction' and 'Data Store' (Level 1). 'Evidence Extraction' receives 'Data Uncertainty (Force / Weight)' and 'Representation Uncertainty'. 'Data Store' receives 'Representation Uncertainty'. Both 'Evidence Extraction' and 'Data Store' feed into 'Fusion' (Level 2). 'Fusion' also receives 'Process Uncertainty'. 'Fusion' outputs to 'Output'. 'Output' is associated with 'Output Uncertainty'. 'Level of Complexity' is an input to the entire system.

The taxonomy of level 2 HLIF types discussed in section 2 defines the complexity of the uncertainty considerations that must be accounted for. Five different types are identified, ranging from simple entity attribute refinement using situation status data to the development of a complete situation assessment assembled from applicable situational fragment data. The uncertainty in the input data / evidence must be assessed for relevance, credibility, and force / weight, per the ontology of evidence presented in Laskey et al. [17]. The representation uncertainties that drive the modeling methodologies can be classified per the uncertainty ontology developed by the W3C Incubator Group for Uncertainty Reasoning [18]. A variety of different models can be used to properly capture the aspects of uncertainty in the data [19, 20]. Finally, the output uncertainty strongly depends on the a priori identification of possible situation choices, or upon having a fusion process that allows for an effective open world assumption. These uncertainty considerations are the beginning of understanding how to evaluate the effectiveness of various uncertainty management methods in high-level fusion.

The Army Research Office provided partial funding to Paulo Costa for this research, under grant W911NF-11-1-0176.

- [1] D. L. Hall, J. Llinas, "Multi-Sensor Data Fusion" in *Handbook of Multisensor Data Fusion: Theory and Practice (2nd ed)*, CRC Press, pp 1-14, 2009
- [2] D. A. Lambert, "A Blueprint for Higher Level Fusion Systems", in *Information Fusion*, pp 6-24, Elsevier, Vol 10, 2009
- [3] M. M. Kokar, C. J. Matheus, K. Baclawski, J. A. Letkowski, M. Hinman, J. Salerno, "Use Cases for Ontologies in Information Fusion" Proceedings of the 7th International Conference on Information Fusion (2004), retrieved from <http://vistology.com/papers/Fusion04-UseCases.pdf> on 1 Jun 2012.
- [4] E. G Little, G. L Rogova, "Designing Ontologies for Higher Level Fusion", *Information Fusion*, pp 70-82, Elsevier, Vol 10, 2009
- [5] P. C. G. Costa (2005) *Bayesian Semantics for the Semantic Web*. Doctoral Thesis, School of Information Technology and Engineering, George Mason University. Fairfax, VA, USA, 2005.

- [6] R. N. Carvalho (2011) *Probabilistic Ontology: Representation and Modeling Methodology*. Doctoral Thesis, School of Information Technology and Engineering, George Mason University. Fairfax, VA, USA, 2011.
- [7] Evaluation of Technologies for Uncertainty Reasoning Working Group (ETURWG) website, <http://eturwg.c4i.gmu.edu/?q=aboutUs>, retrieved May 19, 2012.
- [8] P. C. G. Costa, K. B. Laskey, E. Blasch, A. Josselme, "Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology", Proceedings of the 15th International Conference on Information Fusion (To Be Published)
- [9] A. N. Steinberg, C. L. Bowman, "Revisions to the JDL Data Fusion Model," *Handbook of Multisensor Data Fusion: Theory and Practice (2nd ed)*, CRC Press, pp 45 - 68, 2009
- [10] S. Das, *High-Level Data Fusion*, Boston MA (USA): Artech House, 2008
- [11] J. Sowa, *Knowledge Representation: Logical, Philosophical and Computational Foundations*, Grove CA (USA): Brooks/Cole, Pacific, 2000
- [12] http://www.raytheon.com/capabilities/products/apg63_v3/
- [13] <http://www.p3orion.nl/variants.html>
- [14] Air Power Australia website, <http://www.ausairpower.net/APA-Acquisition-GCI.html#mozTocId55304>, as retrieved on June 2, 2012
- [15] D. A. Schum, *The Evidential Foundations of Probabilistic Reasoning*, New York: John Wiley and Sons, Inc., 1994.
- [16] D. Schum, "Thoughts About a Science of Evidence," University College London Studies of Evidence Science, retrieved from 128.40.111.250/evidence/content/Science.doc on June 2, 2012
- [17] K.B Laskey, D. A. Schum, P. C. G Costa, T. Janssen, "Ontology of Evidence", Proceedings of the Third International Ontology for the Intelligence Community Conference (OIC 2008), December 3-4, 2008
- [18] K.J. Laskey, K. B. Laskey, P. C. G. Costa, M. M. Kokar, T. Martin, T. Lukasiewicz, Uncertainty Reasoning for the World Wide Web, W3C Incubator Group Report 31 March 2008. Retrieved from <http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/>
- [19] P. Walley, "Measures of uncertainty in expert systems", *Artificial Intelligence*, 83(1), May 1996, pp. 1-58
- [20] B. Khaleghi, A. Khamis, F. O. Karray, "Multi-sensor Data Fusion: A Review of the State-of-the-Art", *Information Fusion* (2011), doi: 10.1016/j.inffus.2011.08.001
- [21] K. B. Laskey, P. C. G. Costa, T. Janssen, "Probabilistic Ontologies for Multi-INT Fusion" in Proceedings of the 2010 conference on Ontologies and Semantic Technologies for Intelligence, 2010
- [22] A. N. Steinberg, "Foundations of Situation and Threat Assessment" in *Handbook of Multisensor Data Fusion: Theory and Practice (2nd ed)*, CRC Press, pp 437 -502, 2009
- [23] R. G. Almond, *Graphical Belief Modeling*, New York NY (USA): Chapman and Hall, 1995
- [24] P. P. Shenoy, "Valuation-Based Systems for Bayesian Decision Analysis," *Operations Research*, pp 463 – 484, Vol 40, No 3, May-June 1992,
- [25] P. Svensson, "On Reliability and Trustworthiness of High-Level Fusion Decision Support Systems: Basic Concepts and Possible Formal Methodologies", 9th International Conference on Information Fusion, Florence (Italy), 10-13 July 2006, retrieved online from <http://www.isif.org/fusion/proceedings/fusion06CD/Papers/51.pdf> on May 6, 2012
- [26] A. Karlsson, Dependable and Generic High-Level Algorithms for Information Fusion – Methods and Algorithms for Uncertainty Management, Technical Report HS-IKI-TR-07-003, University of Skovde, retrieved 15 Sep 2011 from his.diva-portal.org/smash/get/diva2:2404/FULLTEXT01.
- [27] E. Franconi, Description Logic Tutorial Course, downloaded from <http://www.inf.unibz.it/~franconi/dl/course/> on 1 May 2012